CESifo Working Papers
www.cesifo.org/wp

Pension-Induced Rigidities in the Labor Market for School Leaders

Cory Koedel
Jason A. Grissom
Shawn Ni
Michael Podgursky

CESIFO WORKING PAPER NO. 3605
CATEGORY 5: ECONOMICS OF EDUCATION
OCTOBER 2011

PRESENTED AT CESIFO AREA CONFERENCE ON ECONOMICS OF EDUCATION, SEPTEMBER 2011

CESifo *Working Paper No. 3605*

Pension-Induced Rigidities in the Labor Market for School Leaders

Abstract

Educators in public schools in the United States are typically enrolled in defined-benefit pension plans, which penalize across-plan mobility. We use administrative data from Missouri to examine how the mobility penalties affect the labor market for school leaders. We show that pension borders greatly affect leadership flows across schools – for two groups of schools separated by a pension border, our estimates indicate that removing the border will increase leadership mobility between them by 97 to 163 percent. We consider the implications of the pension-induced rigidities in the leadership labor market for schools near pension borders in Missouri. Our findings are of general interest given that thousands of public schools operate near pension boundaries nationwide.

JEL-Code: H500, I200, J300.

Cory Koedel
Department of Economics
University of Missouri / USA
koedelc@missouri.edu

Jason A. Grissom
Peabody College of Education and Human Development / Vanderbilt University / USA
jason.grissom@vanderbilt.edu

Shawn Ni
Department of Economics
University of Missouri / USA
NiX@missouri.edu

Michael Podgursky
Department of Economics
University of Missouri
USA – Columbia, MO 65203
PodgurskyM@missouri.edu

September 2011
The authors are faculty at the Peabody College of Education and Human Development, Vanderbilt University (Grissom); and the Department of Economics, University of Missouri (Koedel, Ni, Podgursky). Podgursky is also a Fellow of the George W. Bush Institute at Southern Methodist University. They thank Julie Cullen, Peter Dolton, and Peter Mueser for useful comments and suggestions, and acknowledge research support from the Institute of Education Sciences through the Center for Analysis of Longitudinal Data in Education Research (CALDER). The usual disclaimers apply.

I. Introduction

A burgeoning research literature shows that school leaders play an important role in the education production process (Brewer, 1993; Cannon et al., 2010; Clark et al., 2009; Coelli and Green, 2010; Dhuey and Smith, 2010; Grissom and Loeb, 2011).[1] However, we know relatively little about the labor market for school leaders.[2] One largely unexamined aspect of the leadership labor market is that most school leaders, and prospective school leaders, are heavily invested in defined-benefit (DB) pension plans. A well-known feature of DB plans is that they restrict mobility by penalizing individuals who leave the system prior to retirement eligibility. For most school leaders the mobility penalties are severe. This is because the range over the career cycle when transitions to leadership typically occur coincides with the range when the mobility penalties in DB plans are greatest (Costrell and Podgursky, 2010).

We use administrative panel data from Missouri to examine the effects of pension-system boundaries on leadership mobility. Missouri is an interesting case to study because educators belong to three different pension systems. Roughly 90 percent participate in the state's Public Service Retirement System (PSRS), and the remaining 10 percent are enrolled in one of the district-specific systems in Kansas City (KC) and St. Louis (STL). The presence of the three distinct pension systems in Missouri, along with the availability of statewide administrative data that allow us to track mobility, offers a unique opportunity to evaluate pension-border effects in the leadership labor market.

[1] Improving leadership quality in schools is also a key objective among major actors in education reform. For example, the Knowledge is Power Program (KIPP), a national network of highly successful charter schools from across the country (Angrist et al., 2010), regularly references the important role that school leaders play in the success of the program. The philanthropic community also invests heavily in leadership-based interventions like New Leaders for New Schools, an organization that trains school leaders and is funded by donations from the Gates Foundation, the Broad Foundation, and the Michael and Susan Dell Foundation, among others.
[2] The literature on the leadership labor market is small. Notable recent studies include Cullen and Mazzeo, 2008; Myung et al., forthcoming; and Loeb et al, 2010.

We begin by illustrating the costs associated with crossing a pension border for prospective school leaders, focusing primarily on the transition from teaching to school leadership (teaching is the most common entry path into leadership). Relative to a within-system move from teaching to leadership, an across-system move can result in pension-wealth losses on the order of hundreds of thousands of dollars. Moves from one leadership position to another are similarly costly if a pension border is crossed.

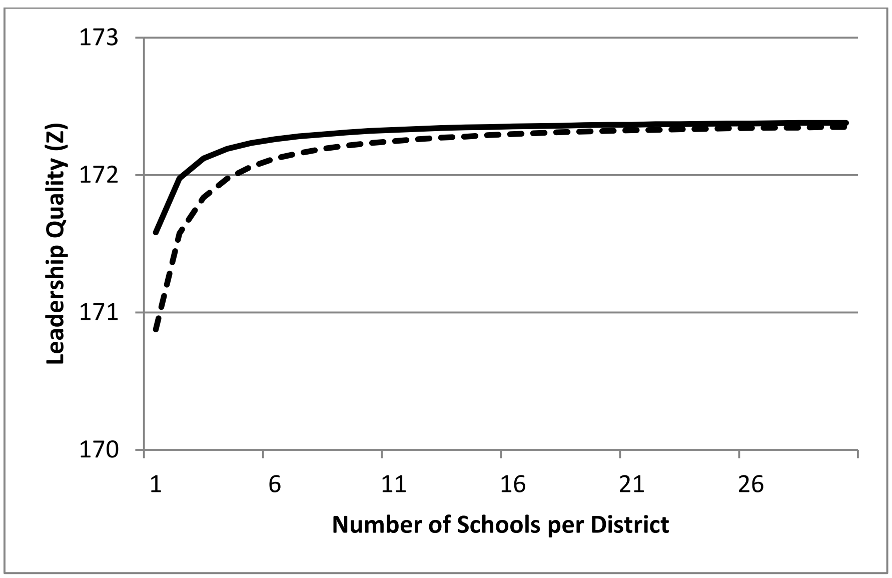
Next we examine data on leadership mobility within and across pension systems in Missouri. The two urban districts in the state, each operating a separate and independent pension system, are dramatically less likely to hire school leaders from an outside district relative to PSRS districts, which all share the same pension system. In fact, the average PSRS district is roughly five times more likely than KC or STL to hire a school leader from the outside. This simple difference suggests that the pension borders affect mobility, but inference is limited because the city districts differ from PSRS districts in many ways. We perform a simulation exercise designed to isolate the pension-border effects in each border region of the state. This exercise shows that higher exposure to pension boundaries significantly lowers interschool labor flows – for two groups of schools separated by a pension border, our estimates indicate that removing the border will increase leadership mobility between them by 97 to 163 percent.

We conclude by examining the consequences of our findings for leadership quality in schools. In Missouri, we use available quality measures to show that the pension borders separate the city school districts from a more-qualified leadership pool on the outside. In a closing discussion we consider the implications of our findings in a larger national context. Across the United States, 22,000 miles of state boundaries, which are also pension boundaries, create rigidities in educator labor markets.

II. Motivation

Basic economic theory suggests that pension-induced rigidities in the leadership labor market will have efficiency implications. Furthermore, they can reinforce or exacerbate inequitable distributions of leadership talent across schools. We briefly illustrate these points in Figures 1 and 2. Both figures show results from a simple, static model of the leadership labor market where two contiguous school districts are separated by an impermeable pension border. The figures illustrate two mechanisms by which pension borders will affect leadership quality in schools: (1) by shrinking the applicant pools from which administrators in border regions can draw, and (2) by reinforcing differences in applicant-pool quality across pension systems.[3]

Figure 1. Average Leadership Quality in Schools from a Simple Two-District Model Where the Distribution of Applicant Quality is the Same on Both Sides of the Pension Border.

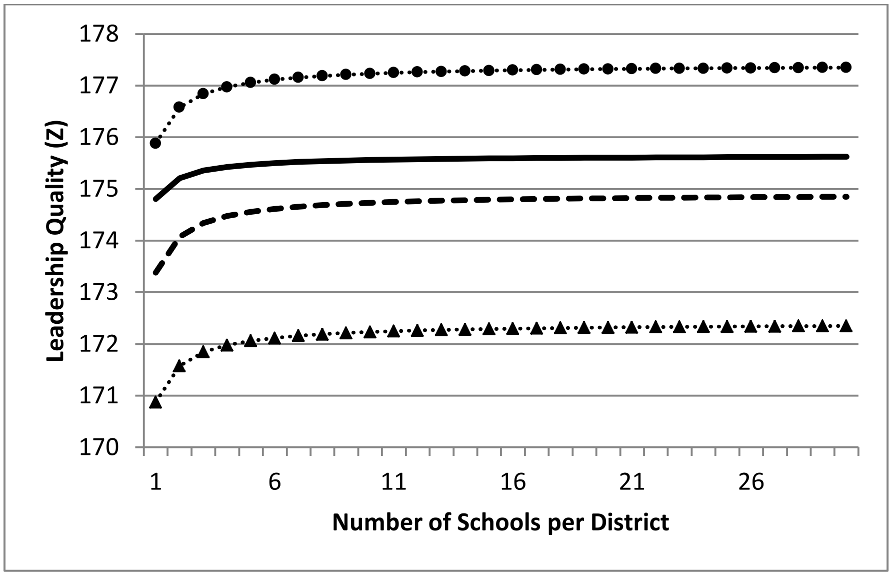


Notes: The dashed line indicates average leadership quality with the pension border; the solid line indicates average leadership quality without the pension border.

[3] The model treats differences in applicant-pool quality across districts as exogenous, but they may not be. For example, in Missouri, because the pension borders restrict mobility for educators in the urban districts, they may affect initial teacher recruitment. Jacob (2007) discusses the general challenges that urban districts face in recruiting teachers – for KC and STL, the restricted mobility is an additional concern. Furthermore, the pension borders may affect the returns to investment in leadership-specific human capital in the different pension systems in Missouri. These issues are beyond the scope of the current project, but merit attention in future research.

Brief model details: District 1 has n_1 schools, and therefore n_1 leadership positions. It has N_1 = (n_1*s) potential candidates, where s indicates the number of candidates per school (i.e., teachers). Similarly, District 2 has n_2 schools and N_2 potential candidates. We set $n_1 = n_2 = n = 1$ and allow the districts to grow up to the point where n = 30.We hold s fixed at 10 throughout. Candidate quality follows the same normal distribution in each district, $Z \sim N(160,50)$.

Figure 2. Average Leadership Quality in Schools from a Simple Two-District Model Where Average Applicant Quality is Higher in District 1



Notes: The dashed line indicates average leadership quality with the pension border; the solid line indicates average leadership quality without the pension border; circles indicate average leadership quality at District 1 (stronger applicant pool); triangles indicate average leadership quality at District 2 (weaker applicant pool).

Brief model details: Same as in Figure 1 except that applicants in District 1 come from a new distribution: $Z_1 \sim N(165,50)$. Applicants from District 2 continue to come from the original distribution: $Z_2 \sim N(160,50)$.

In the model, administrators fill leadership vacancies by choosing the highest-quality candidate from the pool of applicants, where quality is perfectly observed and measured by Z. The two districts are equal-sized, and we allow them to grow together from containing 1 to 30 schools. Figure 1 shows average leadership quality in schools, with and without the pension border, when applicants are drawn from the same underlying quality distribution in both districts

($Z \sim N(160,50)$).[4] When the pension border separates small districts the reduction in leadership quality is substantial – a result that is relevant for rural districts near state lines. Consider the case of two single-school districts for intuition: pooling the districts raises average quality overall when the top candidate in one district is weaker than the second-best candidate in the other. For larger districts, Figure 1 shows that the border effect on the size of the applicant pools alone is of little practical significance.

In Figure 2 we adjust the underlying distribution of applicant quality in District 1 so that it has a stronger pool ($Z_1 \sim N(165,50)$; $Z_2 \sim N(160,50)$). Unsurprisingly, even as the districts become large, overall leadership quality is lower with the pension border. This is because some applicants in District 1 who are not hired into leadership positions have higher values of Z than some leaders in District 2. The effect of the border on leadership quality overall is increasing with the gap between $\bar{Z}_1$ and $\bar{Z}_2$. An additional concern is that the border will reinforce and potentially exacerbate inequality if District 2 is more disadvantaged in other ways.

Our simple, static model clearly oversimplifies the leadership labor market. Nonetheless, it is useful for illustrating the potential efficiency and equity consequences associated with pension borders. Pension borders will also reduce the quality of leader/school matches in border regions. Using data on teachers from North Carolina, Jackson (2010) shows that teacher/school match quality explains a large fraction of what we commonly describe as teaching effectiveness (measured by value-added). If leader/school match quality is similarly important, then lower-

[4] The distribution of Z was selected because it roughly mimics the distribution of one of the observable measures of leadership quality that we consider in Section VII (licensing exam scores), although the specifics of the distribution are not crucial to gain inference from the figures.

quality matches owing to the pension-induced mobility restrictions are an additional consequence of pension boundaries.[5]

III. Background

All three pension systems in Missouri, like nearly all educator pension systems nationwide, are defined-benefit plans. The general formula used to compute the annual payment in each system is:

$$Y_i = F_i * YOS_i * FAS_i \qquad (1)$$

In (1), Y_i is the annual pension payment for individual *i* at the time of retirement. The payment depends on a formula factor, F_i; years of system service, YOS_i; and the individual's final average salary, FAS_i, which is calculated based on the average of the highest few years of system earnings. Table 1 reports the parameters for each system in Missouri in the year 2000 (Appendix Table C.1 shows changes to the system rules over the course of our data panel). While some specific features differ across the systems, their general structures are similar.

Table 1. Pension Plan Parameters in 2000

	PSRS	KC	STL
Vesting (Years)	5	5	5
Social Security	No	Yes	Yes
Retirement Eligibility (Normal and Early)	Full Retirement: Age 60 or 30 YOS; Early Retirement: Age 55, Rule of 80, 25-and-out	Full Retirement: Age 60; Rule of 75 Early Retirement: Any Age with YOS ≥30	Full Retirement: Age 65; Rule of 85 Early Retirement: Age 60
Formula Factor at Full Retirement	0.025	0.020, *F*YOS* capped at 0.60	0.020, *F*YOS* capped at 0.60
FAS	Highest three years	Highest four years	Highest three years (in last 10 years)
COLA	Yes (80% Cap)	No	No

[5] Of course, the rigidities that we document in the leadership labor market will also affect teachers, in which case Jackson's (2010) findings are of direct relevance.

A key difference between the systems is that educators in the city systems participate in Social Security, while educators in PSRS do not.[6] This difference facilitates our analysis; without it there would likely be reciprocity between the pension systems. For example, in other city districts that operate their own pension systems (e.g., New York, Chicago), educators can move between the state and local systems without incurring large pension-wealth losses. But reciprocity is more challenging in Missouri because system benefits and participant contributions are very different across the systems that do and do not incorporate Social Security (see Table 1). Thus, the three pension systems within Missouri are separate, and differ in the same way that systems differ across states.[7]

Two common features of the pension systems penalize mobility. First, like other educator pension systems nationwide (Costrell and Podgursky, 2009), all three systems offer generous retirement provisions that depend on within-system experience, *YOS*. These provisions include the rules of 80, 75 and 85 in PSRS, KC and STL, respectively. The rule amounts refer to individuals' combinations of age and system experience – once the sum of these two numbers reaches the level of the rule, the member can collect full retirement benefits. For example, although the official (and in earlier decades, typical) retirement age in KC is 60, an age-50 teacher with 25 years of system service is eligible to retire without penalty and begin collecting benefits immediately. PSRS also has an additional provision called "25-and-out," which stipulates that an individual can exit and begin collecting benefits immediately, regardless of age, so long as he or she has 25 years of system experience (with a penalty that is much less than is

[6] State and local workers were originally excluded from the Social Security system; however, Congress passed legislation in the early 1950s that permitted states to include their employees. Most states chose to include their teachers in Social Security, though currently about 30 percent are not covered. Teachers in most states are either in or out of Social Security. Missouri is somewhat unusual in this regard (Munnell, 2000).

[7] Although benefit levels and contributions differ across the systems because of the Social Security difference, the structural components that cause the rigidities in the labor market are unaffected. See below for details.

actuarially appropriate). In all three systems, and in most other educator pension systems, if an individual exits the pension system too soon she loses the option to exit under one of these provisions. As we illustrate below, this is very costly.

The second common feature of the systems that penalizes mobility – and again, a feature common to DB systems generally – is that *FAS* is frozen at the time of exit. To illustrate why this is important, consider two individuals who end up with the same wage profile over the course of their respective 30 year careers. The first individual stays in the same pension system and her final payment is equal to $30*F*FAS$, where *FAS* is calculated based on her last few years of earnings. The second individual switches systems after 15 years. Her final payment comes from the two systems and is equal to $\{15*F*FAS_1 + 15*F*FAS_2\}$, where FAS_1 is based on her final average salary at the time of her exit from the first system, unadjusted for inflation or life-cycle pay increases.[8]

We illustrate the mobility costs associated with crossing pension-system boundaries by comparing pension-wealth accrual under various career scenarios for a representative teacher in Missouri. The key comparisons are between within- and across-system moves from teaching to school leadership. We do not explicitly document the costs associated with moves from one leadership position to another, but they are substantively very similar.[9]

Pension wealth is calculated as the present discounted value of the stream of pension payments. An individual's pension wealth at time s, with collection starting at time j, where $j \geq s$, can be written as:

[8] Where they exist, COLA's in educator pension systems apply to the pension annuity at retirement, not to the *FAS* for those who terminate early. COLA's are explicit in PSRS, and implemented on an *ad hoc* basis in KC and STL.
[9] More than anything else, the pecuniary costs associated with crossing a pension-system border depend on the individual's combination of age and experience at the time of the move. Therefore, regardless of whether we choose to illustrate mobility costs for teacher-to-leader or leader-to-leader moves, the costs will be very similar for individuals who move with similar age/experience profiles.

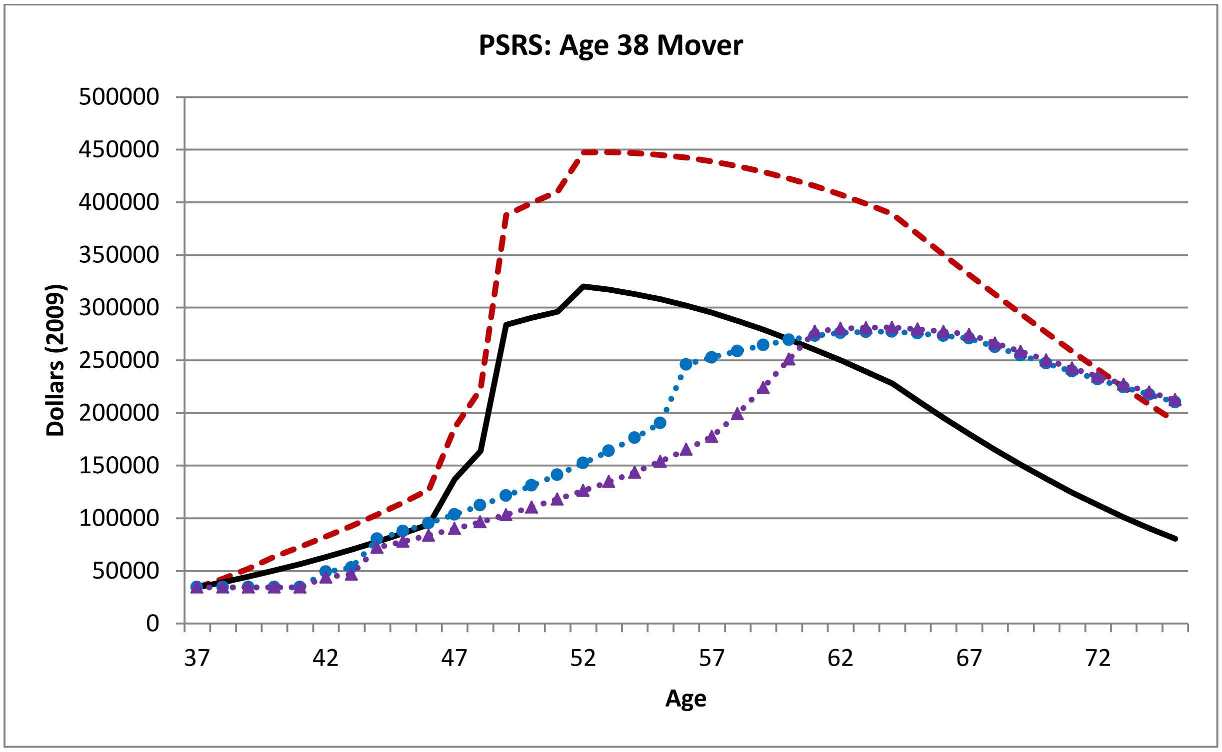
$$\sum_{t=j}^{T} Y_t * P_{t|s} * d^{t-s} \tag{2}$$

In (2), Y_t is the annual pension payment in period t, $P_{t|s}$ is the probability that the individual is alive in period t conditional on being alive in period s, and d is a discount factor.[10] In KC and STL, the pension payment is the sum of the system and Social Security payments. We allow the individual to begin collecting payments from each source at the optimal time. Additional details about our pension-wealth calculations are provided in Appendix A.

Figure 3 shows pension-wealth accrual over the career cycle for a representative teacher who is currently 37 years old and began teaching in PSRS at the age of 25. Her pension-wealth profile is illustrated under several different career circumstances moving forward. First, the solid black line shows pension-wealth accrual under the scenario where she remains in teaching within PSRS permanently. The uneven accrual rate displayed in the figure has been well-documented (see, for example, Costrell and Podgursky, 2009).

[10] We set $T = 101$ for our calculations. After accounting for the impact of discounting and the survival probabilities, the calculations are qualitatively insensitive to reasonable adjustments to this threshold.

Figure 3. Illustration of the pension-wealth profiles for an age-38 teacher who (1) remains in teaching in PSRS for her entire career (solid black line), (2) moves from teaching to leadership within PSRS at the age of 38 (red dashed line), (3) moves from teaching to leadership at the age of 38, but moves from PSRS to KC in the process (blue dotted line, circles), (4) moves from teaching to leadership at the age of 38, but moves from PSRS to STL in the process (purple dotted line, triangles). Based on year-2000 pension rules in each system.



The remaining pension-wealth profiles correspond to moves from teaching to school leadership at age-38.[11] The dashed red line shows pension-wealth accrual when the move occurs within PSRS. The within-system move corresponds to a large increase in pension wealth (driven by the increase in salary that comes with the transition – see Appendix A for details). The other two lines show pension-wealth accrual in scenarios where the move involves crossing a pension-system boundary; the lines marked by circles and triangles are for moves to KC and STL,

[11] Age-38 is the median age-at-entry into leadership in the Missouri data. The distribution of age-at-entry into leadership shifts upward slightly if we omit the rural districts in the state, but the adjustment is qualitatively insignificant. Appendix Table C.2 shows calculations for moves at different points in the career cycle.

respectively. In both border-crossing scenarios the move corresponds to a dramatic reduction in pension wealth. For example, movement to a leadership position in KC from PSRS produces a 63 percent loss in pension wealth when compared to a similar leadership promotion within PSRS (-283,638 in 2009 dollars), holding the retirement age constant at the peak-value of pension wealth in PSRS.[12] If we amortize this over the remainder of her career it amounts to an annual loss of roughly $25,500.

The pension losses illustrated in Figure 3 are not unique to the move scenarios considered in the figure. In Appendix Table C.2 we show that moves across any pension border, and over most of the range of the career cycle when leadership work typically occurs, are associated with large losses in pension wealth.[13] The extent to which labor flows are actually restricted by these mobility penalties is an empirical question, to which we now turn our attention.

IV. Data

We use an 18-year administrative data panel from the entire state of Missouri for our analysis. Data are available from the 1992-1993 through 2009-2010 school years, and include information about each public school employee's position type, salary, plus information on the schools and districts in which the person worked. The length and breadth of our data panel is a key feature of our study as it ensures that we observe a sufficient number of leadership moves.

For our primary analysis we group principals and assistant principals together, and refer to them synonymously as "school leaders."[14] We also report some results separately for full

[12] Most educators retire fairly close to the peak-value year (the year when pension-wealth is maximized), although there is some variability. See Coile and Gruber (2007) for a discussion of the peak-value construct and alternatives (also see Stock and Wise (1990a,b)).
[13] The exception is that more-senior school leaders can marginally benefit in some cases by switching pension systems to double-dip; however, in Appendix Table C.3 we show that double-dipping does not explain the limited cross-border mobility that we observe in the Missouri data.
[14] Correspondingly, for the construction of the wage profiles in our pension-wealth calculations from the previous section, we combine moves into principal and assistant principal positions. So, for example, when we report that the average immediate salary increase corresponding to a move from teaching to leadership is 30 percent, this

principals. In total, we observe 11,034 leadership hires over the course of the data panel, where we define a leadership hire to have occurred whenever a leader starts in a new school. Nearly 50 percent of new hires (5,508) represent first-time leaders, and of those, 67 percent come directly from teaching in Missouri schools. The remainder come mostly from other school- and district-level supervisory positions (e.g., special education administrators, program coordinators) and guidance counselor positions. Table 2 provides basic summary statistics for the data.

Table 2. Descriptive Statistics

	N		Percent of Total Hires
Total School Leadership Hires	11034		
Principal Hires	6313		57.2%
Assistant Principal Hires	4721		42.8%
First-Time School Leadership Hires (In-State)	5508		49.9%
From Teaching Positions		3706	33.6%
From Supervisor/Other Administrative Positions		1004	9.1%
From Guidance Counselor		198	1.8%
From Other		137	1.2%
Prior Position Not Recorded		463	4.2%
Moves from Other Leadership Positions (In-State)	5268		47.7%
From Principal Positions		2255	20.4%
From Assistant Principal Positions		2513	22.8%
From Other		500	4.5%
Moves from Out-of-State	258		2.3%

Characteristics of School Leader Hires	Mean	St. Dev
Female	0.47	0.50
Nonwhite	0.15	0.35
Salary (in $2009)	68,307	15,485

Table 3 shows within- and out-of-district hire rates for school leaders in KC, STL, PSRS, and three subsamples of PSRS districts (out-of-state hires are included with out-of-district hires).

calculation is based on all moves from teaching into either full or assistant principal positions. None of our findings are substantively sensitive to how we define "school leader."

The PSRS subsamples were selected to be comparable to the city districts along a particular dimension: column (4) shows hire rates for PSRS districts that are geographically close to the city districts (neighbors), column (5) shows hire rates for districts that are similarly disadvantaged, and column (6) shows hire rates for the ten largest districts in PSRS.[15]

Table 3. Within- and Out-of-District Leadership Hires for PSRS, KC, STL, and Subsamples of PSRS Districts

	PSRS All	KC	STL	PSRS Neighbors	PSRS Disadvantaged	PSRS Large Districts
<u>All Leadership Hires</u>						
Within-District Hires	54.5	90.4	90.0	61.0	58.7	72.8
Out-of-District Hires	43.7	8.5	9.3	37.8	40.1	26.2
Unknown (in-state)	1.8	1.1	0.8	1.2	1.2	1.0
N	9668	707	659	2214	414	1818
Avg. % Free/Reduced Lunch	31.7	59.9	70.8	27.6	60.3	13.5
Avg. % Disadv. Minority	7.1	70.9	83.3	38.9	79.3	18.2
Average Enrollment	1495	32839	39948	6708	3121	17948
Number of Districts	539	1	1	32	9	10

Notes: PSRS Neighbors include districts that are geographically adjacent to either KC or STL, or are adjacent to the adjacent districts. The neighboring districts are within a commutable distance to the city in each region. Out-of-state hires are included with out-of-district hires. Unknown in-state hires are individuals with previous Missouri experience but for whom we cannot identify their place of prior employment.

Because KC and STL operate their own pension systems, all out-of-district leadership hires necessarily cross a pension border. Districts within PSRS can share hires without the occurrence of a border crossing.[16] Consistent with the mobility penalties illustrated above, the raw data show that the city districts have much lower out-of-district hire rates than their PSRS counterparts. In fact, the out-of-district hire rate for the typical PSRS district is roughly five times that of either of the city districts. But it is difficult to gain inference about the pension-border effects from Table 3 because the city districts differ from PSRS districts in many ways,

[15] The neighboring PSRS districts are well within a commutable distance to the city districts. They include districts that touch the city districts, or touch the districts that touch the city districts. In each region, the neighboring districts cover a much smaller area than the Labor Market Area as defined by the Bureau of Labor Statistics (BLS, 2011).
[16] Roughly 3 percent of all out-of-district hires in PSRS districts come from KC or STL. For brevity, we do not separate these moves in the tables. Appendix Table C.3 also shows within- and out-of-district hire rates for several subgroups of school leaders – the patterns in Table 3 are replicated regardless of how we divide the data.

even when we consider the PSRS subsamples (see the bottom panel of the table).[17] Ultimately, while the raw data are consistent with the pension borders reducing mobility, inference about the border effects is clouded by the general lack of comparability between the schools across pension systems.

V. Empirical Strategy

We take two approaches to isolate the pension-border effects on leadership mobility. First, in our preferred analysis, we randomly regroup schools in the border regions into simulated "artificial districts," and use chance variation from the re-districting procedure to identify the border effects. Second, we estimate simple, pairwise models that measure the influence of the pension borders on leadership sharing between pairs of schools in each border region.

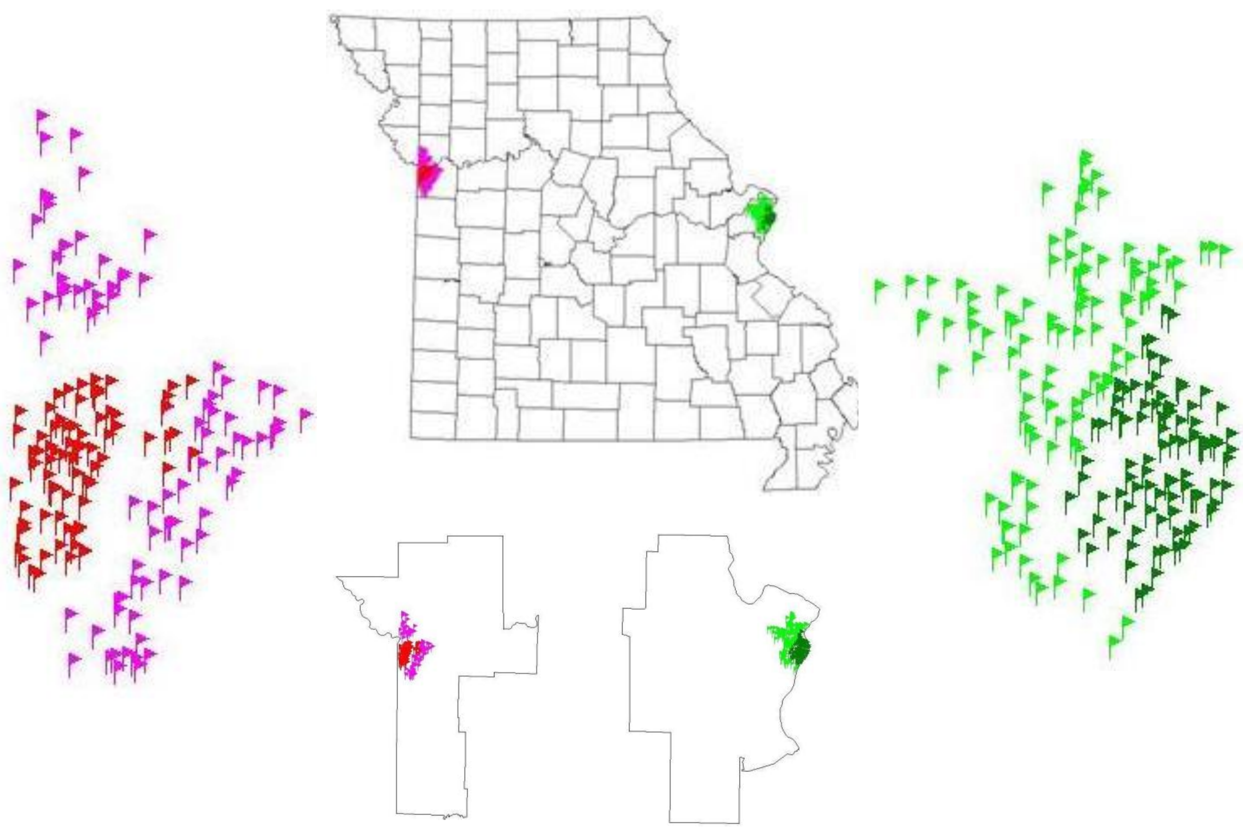
Artificial District Simulations

It is useful to consider momentarily how an experiment could be constructed to isolate the pension-border effects. Ideally schools could be randomly assigned to different pension systems (by either moving the borders or schools at random), in which case it would be straightforward to determine whether labor flows are restricted by the borders. But this type of random assignment is clearly infeasible. As a next-best alternative, we perform a simulation exercise where we randomly re-group schools from the border regions into artificial districts that are differentially segmented by the real pension borders. We then use chance variation in the degree of segmentation of the artificial districts to identify the pension-border effects. We show that the simulations can be designed so that the border effects are identified separately from other factors that are likely to influence mobility.

[17] The neighboring districts are too small and not disadvantaged enough, the disadvantaged districts are too small, and the largest PSRS districts are neither disadvantaged enough nor big enough. Additionally, some of the largest non-city districts in Missouri are fairly isolated geographically, which likely reduces their out-of-district hire rates.

The first step in our artificial re-districting procedure is to identify two fixed samples of schools from districts that surround the borders. In each region, we use the city schools and a subsample of the PSRS neighbors from Table 3 (the reason that we do not use all neighboring schools will become clear momentarily). The locations of the schools in each simulation sample are illustrated in Figure 4, where the darker flags are city schools and the lighter flags are PSRS schools.

Figure 4. Illustration of Data Samples for Artificial-District Simulations. Left: Schools Used in KC-Area Simulations. Center Top: Zoomed-Out View of KC- and STL-Area Schools. Center Bottom: KC- and STL-Area Schools Relative to the BLS Labor Market Areas. Right: Schools Used in STL-Area Simulations. The Darker Flags in Each Region Correspond to City Schools.



The schools in the neighboring PSRS districts in Table 3 are within a commutable distance to the city, and the subsamples of neighboring schools that we use for the simulations are even closer. This is illustrated in the bottom-middle panel of the figure, which overlays the Labor Market Area in each region as defined by the Bureau of Labor Statistics (BLS, 2011). The BLS describes a Labor Market Area as “an economically integrated geographic area within which individuals can reside and find employment within a reasonable distance or can readily change employment without changing their place of residence.” The samples of schools in each region are comfortably within the Labor Market Area.[18]

For the simulations, we randomly re-assign the schools in each region into ten artificial districts that consist of an equal number of schools, irrespective of schools’ actual district assignments, and repeat the reassignment process 1,000 times.[19] At each iteration we capture several pieces of information about the artificial districts. First, we calculate the geographic proximity of the schools within each artificial district, measured as the average distance between them. Next, we measure each artificial district’s basic demographic characteristics including district enrollment, minority share, and free/reduced-lunch share.[20] We use the information about the minority and free/reduced-lunch shares to measure the similarity between the schools. Consider the minority-share variable as an example. As the average minority share across schools within the artificial district approaches 0.50, within-district diversity increases; as it approaches zero or one, diversity decreases.[21] We construct the similarity measures as the absolute value of the difference between the across-school average of the given characteristic in

[18] Also note that both Labor Market Areas in Missouri are connected to other Labor Market Areas on the other side of the state border. In fact, one could argue that the Labor Market Areas are too big. But even if we cut them in half, our samples of schools still fit comfortably within them.

[19] More precisely, we create districts with as close to an equal number of schools as possible. The first nine artificial districts have the same number of schools and the tenth district is slightly larger or smaller than the others. This is an important feature of the simulation design; in unreported results we find that within PSRS, district size is the most important determinant of the within-district hire rate.

[20] The districts are constructed to include the same numbers of schools, but there is still variation in total enrollment.

[21] This measure of similarity would not be useful in a world where all schools are similarly diverse (e.g., if every school had a 50-percent minority share). However, in the data, there is considerable heterogeneity across schools in their shares of disadvantaged students.

artificial district *j*, and 0.50. The larger the difference, the more similar are the schools within the artificial district.

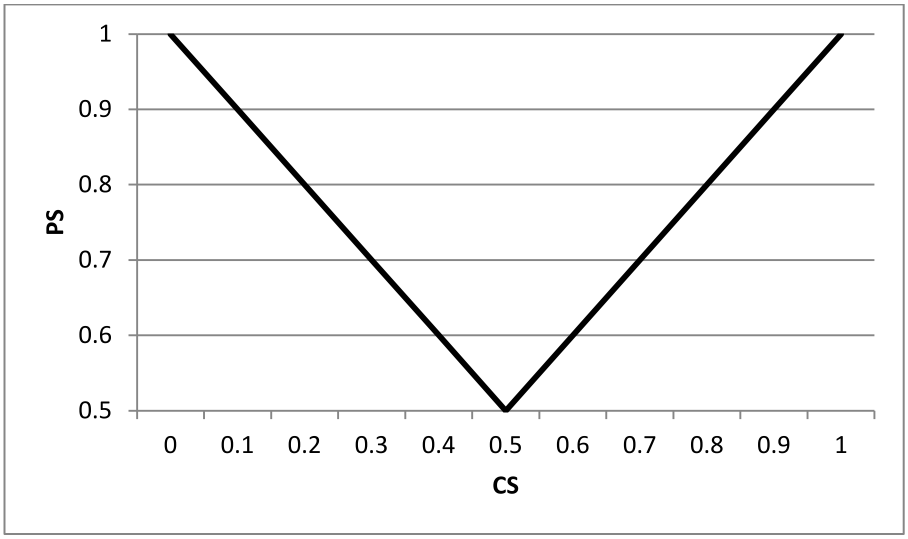
Next we measure the amount of real pension-system sharing within the artificial districts. Given that schools in each region belong to one of two real pension systems, pension sharing can be measured by $PS = max\{CS, (1 - CS)\}$, where *CS* is the share of city schools. The variable *PS* indicates the degree to which each artificial district is segmented by the actual pension border. Larger values indicate more pension-system sharing among schools within the artificial district; smaller values indicate less pension-system sharing.[22]

The function that maps values of *CS* to values of *PS* is V-shaped, with the minimum value of *PS* occurring when $CS = 0.50$. This is illustrated in Figure 5. The correlation between *PS* and *CS* can be broken in the simulated data by constructing the samples of schools so that $E[CS] = 0.50$ in each artificial district. With this in mind we selected our initial samples of schools in the KC and STL regions so that there are equal numbers of city and non-city schools. This allows us to identify the pension-border effects free from other confounding factors associated with the city districts (including the city-share of schools itself).[23]

[22] More pension sharing within the artificial district opens up more unobstructed conduits for mobility between the schools in that district.

[23] As noted above, we mostly selected entire districts to get the PSRS schools in Figure 4. After selecting groups of districts so that the city share was nearly half in each region, we then trimmed a handful of randomly selected schools from each dataset to equate the shares of city and PSRS schools.

Figure 5. Function Mapping City Share (CS) to Pension Sharing (PS)Within Artificial Districts from the Simulations



The final piece of information that we record for each artificial district is the number of leadership hires that are shared between schools. Our hypothesis is that artificial districts that are more segmented by the real pension border – that is, where PS is closer to 0.5 – will have fewer intra-district hires.[24]

We analyze the simulated data using the following regression model, estimated separately for the KC and STL regions:

$$WC_j = \delta_0 + X_j\delta_1 + D_j\delta_2 + CS_j\delta_3 + PS_j\delta_4 + \varepsilon_j \quad (3)$$

In (3), WC_j measures the number of within-district hires for artificial-district j, measured as a count. X_j includes controls for district size and the similarity of the schools within the artificial district, and D_j measures the average geographic distance between the schools.[25] CS_j indicates the share of the artificial district that is comprised of city schools, for which the effect can be

[24] We do not count same-school hires because they cannot be affected by the differences in pension-system sharing across the simulations. Therefore, we estimate the effects of pension-system sharing on across-school labor flows.

[25] Given the close geographic proximity of the schools in each region (Figure 4), distance should not play an important role in our analysis. To be more concrete, geographic proximity will affect the labor market nonlinearly, and for small changes in distance we expect the effect to be essentially zero. The closeness of the schools in our simulation samples suggests the proximity effect should be small. The results in Table 6 confirm this intuition.

separately identified per the above discussion (see Figure 5).[26] PS_j measures pension sharing and is the independent variable of interest.

Obtaining the correct standard errors for the coefficients in (3) is not straightforward. The simulation dataset in each region includes data from 10,000 artificial districts generated in 1,000 batches of 10, and the error terms from equation (3) are correlated both within and across batches. The within-batch correlations can be addressed using standard clustering techniques. However, there is no standard procedure to account for across-batch correlation in the error terms, which is generated by unstructured data overlap across simulations (each batch of 10 schools is constructed by re-dividing schools from the same underlying pool).[27]

On the one hand, the probability that there are two identical artificial districts in the simulated data in either region is negligible. Nonetheless, there will be overlapping groupings of schools across batches throughout, meaning that all of the data across all of the batches can be correlated. The typical clustering solution, which imposes a block-diagonal structure on the variance-covariance matrix, cannot be applied.

In Appendix B we provide an alternative solution to address the issue of data overlap. We develop a procedure by which the *effective number of simulations* in each regional analysis can be calculated, where the effective number is always less than the actual number (1,000). The intuition is that each batch of artificial districts provides some new identifying variation for our analysis, but it also contains repeat information due to the overlapping data.[28] The procedure

[26] We also estimate models where we replace CS_j with a Herfindahl index for each artificial district that indicates its stratification across actual district lines, and obtain substantively similar results. We cannot include the Herfindahl indices and city shares at the same time because they are highly correlated (in the KC region the correlation is roughly 0.94; in STL it is 0.98).
[27] There are 184 schools in the KC region and 266 in the STL region.
[28] Re-dividing of the data again and again produces new variation with which to identify the border effects up to the point where all possible combinations of schools have been exhausted. The marginal new information gained will

outlined in the appendix allows us to account for the fact that our effective sample sizes are smaller than would be the case if each batch of districts were independently drawn. It can be applied to the general case where clustering in the data is unstructured and/or of unknown form. In the present application, we calculate that the effective number of simulations is 350 in the KC region, and 410 in the STL region. We adjust the standard errors from our simulation regressions to account for the data overlap.

Pairwise Models

As an alternative to the simulations we also perform a pairwise analysis using the universe of pairs of schools in each regional dataset. In the KC region there are 16,836 unique pairs of schools; in STL there are 35,245 pairs.

A limitation of the pairwise analysis is that most school pairs do not share any leadership hires, so the outcome variable is mostly populated with zeros. The high prevalence of zeros makes linear regression unappealing. Therefore, we convert our measure of leadership sharing to a simple binary indicator, and estimate a logit model that predicts the occurrence of any shared hire between each pair of schools. In doing so, we throw out information in the data for pairs that share more than one hire.[29] We estimate the following model with the pairwise data:

$$AS_j = \gamma_0 + X_j\gamma_1 + D_j\gamma_2 + DL_j\gamma_3 + PL_j\gamma_4 + u_j \tag{4}$$

In (4), AS_j is an indicator equal to one if any shared hire is observed between the schools in pair j, X_j includes the similarity measures for the schools in the pair and their combined enrollment, and D_j measures the distance between the schools. There is more variation in distance in the pairwise models and unlike in the simulation analysis, distance is important. We include linear

decrease as the number of re-divisions approaches the number of possible combinations, although in our analysis we are far away from this threshold.

[29] In the KC and STL regions respectively, 26 and 18 percent of all pairs of schools with a leadership hire have more than one hire over the sample period. These are converted to unitary values in the logit model.

and quadratic distance terms in the regressions because both are generally significant. DL_j and PL_j are indicators for the schools being divided by a district line and a pension line, respectively. Pension-line crossings always involve crossing a district line, but the reverse is not true.

The model in (4) is similar to the model in (3). The key benefits of the pairwise model are (1) no re-sampling is required because the data include the exhaustive set of pairs and (2) we can directly and separately control for district-line crossings.[30] Again, the drawback is that the pairwise model ignores information about multiple shared hires between schools.

VI. Results

Simulation Analysis

In the real Missouri data, crossing a pension boundary means moving between dissimilar school districts (see Table 3). The key benefit of the simulation design is that the artificial districts do not differ systematically along key dimensions, while at the same time they are differentially segmented by the real pension borders. Table 4 confirms this feature of the simulated data. It shows that pension sharing in the artificial districts is essentially uncorrelated with the minority share, free/reduced-lunch share, enrollment level, and the share of city schools.

Table 4. The Range of *PS* (Pension Sharing) in the Simulated Data, and Correlations Between Artificial-District Characteristics and *PS*

	KC	STL
Range of *PS* (pension sharing)	0.50 – 0.94	0.52 – 0.81
Correlation between *PS* and:		
Share eligible for free/reduced lunch	0.012	0.008
Share nonwhite	-0.007	-0.003
Enrollment (log)	-0.023*	-0.013
Share of city schools	0.006	0.005

** Indicates statistical significance at the 1 percent level.

[30] As noted in footnote 26, our findings in the simulation analysis are qualitatively similar if we replace the city-share control with a Herfindahl index for real-district diversity within artificial districts (the Herfindahl index and the city share variables are too highly correlated to be included simultaneously in the models). The Herfindahl-index control in equation (3) is analogous to the district-line control in equation (4), but is a less-direct measure.

* Indicates statistical significance at the 5 percent level.
Notes: The larger the pension-sharing variable, the less segmented is the artificial district.

But the fact remains that the schools and borders are in geographically fixed locations, and there are differences in the student populations on different sides of the border in each region. This can be seen in Table 5, where we show that pension sharing *is* correlated with the similarity measures. The correlations in Table 5 reflect the fact artificial districts with more schools on either side of the border will have student populations with more in common. The actual differences between schools on different sides of the border are larger in the KC region, and correspondingly, so are the correlations between the similarity measures and pension sharing. Also, of course, the average distance between schools decreases with pension sharing given that the real pension boundaries are geographically defined.[31] The model in (4) explicitly controls for the similarity measures and the geographic proximity of the schools within each artificial district to ensure that these factors are not driving our findings.

Table 5. Correlations in the Simulated Data Between the Artificial-District Similarity and Distance Measures, and *PS* (Pension Sharing)

	KC	STL
Correlation between *PS* and:		
Free/reduced-lunch similarity	0.350**	0.073**
Racial similarity	0.300**	0.003
Average distance between schools	-0.084**	-0.114**

** Indicates statistical significance at the 1 percent level.
* Indicates statistical significance at the 5 percent level.
Notes: The larger the pension-sharing variable, the less segmented is the artificial district.

Table 6 presents our results from the regressions separately for the KC and STL regions. The models are increasingly detailed moving from left to right in the table. The final column shows results from the complete model, but where we track moves for full principals only. All of

[31] The correlation between pension sharing and distance is weaker than for the other similarity measures. This is because of the inner/outer ring construction of the simulated data. Pension sharing in the outer ring is not strongly correlated with the geographic proximity of schools.

the control variables in the models are important predictors of labor-market connectivity.[32] But the key finding from Table 6 is that regardless of which other features of the artificial districts are included in the model, pension sharing remains a statistically significant and economically meaningful determinant of interschool leadership sharing.

Table 6. Regression Results for Artificial-District Simulations. Dependent Variable: Shared Leadership Hires

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5 Full Principals Only
Kansas City						
Pension sharing (PS)	**6.08**	**5.97**	**4.08**	**6.10**	**6.21**	**1.85**
	(1.21)**	**(1.22)****	**(1.24)****	**(1.24)****	**(1.24)****	**(0.59)****
KC schools share		12.56		12.51	12.87	4.33
		(0.71)**		(1.02)**	(1.12)**	(0.37)**
Free/reduced-lunch similarity			-13.10	-3.34	-3.53	1.03
			(2.13)**	(2.16)**	(2.18)**	(1.17)
Racial similarity (nonwhite)			20.40	4.17	4.20	0.42
			(1.51)**	(1.89)*	(1.94)**	(1.00)
Enrollment (log)			9.71	9.88	9.87	1.38
			(0.48)**	(0.47)**	(0.47)**	(0.22)**
Average distance (miles)					0.07	0.12
					(0.10)	(0.05)*
St. Louis						
Pension sharing (PS)	**4.41**	**4.37**	**4.25**	**4.45**	**4.17**	**2.24**
	(1.10)**	**(1.09)****	**(1.04)****	**(1.02)****	**(1.03)****	**(0.62)****
STL schools share		4.77		5.56	5.14	0.21
		(0.64)**		(0.69)**	(0.70)**	(0.46)
Free/reduced-lunch similarity			7.61	3.77	4.08	5.38
			(1.68)**	(1.69)**	(1.69)**	(1.08)**
Racial similarity (nonwhite)			8.52	6.67	5.30	3.92
			(1.47)**	(1.48)**	(1.65)**	(1.04)**
Enrollment (log)			9.91	10.38	10.43	2.84
			(0.38)**	(0.39)**	(0.39)**	(0.24)**
Average distance (miles)					-0.21	-0.11
					(0.11)*	(0.07)

** indicates statistical significance at the 1 percent level.
* indicates statistical significance at the 5 percent level.
Notes: Robust standard errors are in parentheses, clustered within simulations and adjusted for the effective number of simulations as described in Appendix B. The "principals only" model excludes assistant principals.

[32] In the KC results (all leaders), it is seemingly counterintuitive that within district similarity in the share of students who are eligible for free/reduced lunch enters negatively into the equation. However, when all of the similarity controls are considered simultaneously (including the city share of schools), within-district similarity is positively associated with leadership sharing.

How large are the pension border effects that are implied by our estimates? We answer this question by extrapolating from the simulated data to compare the cases where $PS_j = 0$ and $PS_j = 1$. The first case is mechanically ruled out by the construction of our simulations, but can be conceptualized as a polar case where each school in the artificial district has its own DB system. Thus, there is no pension sharing, and all interschool labor flows represent pension-border crossings.[33] At the other extreme is the case where $PS_j = 1$, where there are no pension borders and all schools share the same DB system. The pension-border effect on mobility is the difference in leadership flows between these two cases.

In each region we use the regression coefficients from the full model to compare the predicted number of shared leadership hires when $PS_j = 0$ to the predicted number when $PS_j = 1$. We hold all non-pension variables fixed at their sample averages.[34] Our calculations indicate that going from complete segmentation to total integration increases leadership flows between schools by 97 (in STL) to 163 (in KC) percent. In the full-principal models, the simulations suggest an analogous increase in leadership flows of 98 to 113 percent.[35]

Pairwise Analysis

Table 7 reports marginal effects from the pairwise logit models. For brevity we report estimates from the fully-specified models only, and for ease of presentation the coefficients and standard errors are multiplied by 100. The pension-border effects in Table 7 are consistent in

[33] In a region with one real pension border this case would be equivalent to measuring labor flows only between schools on different sides of the border (ignoring within-system moves).
[34] That is, for each region we use our estimates from the full model to compare $(\hat{\delta}_0 + \bar{X}_j\hat{\delta}_1 + \bar{D}_j\hat{\delta}_2 + \overline{CS}_j\hat{\delta}_3 + 0*\hat{\delta}_4)$ and $(\hat{\delta}_0 + \bar{X}_j\hat{\delta}_1 + \bar{D}_j\hat{\delta}_2 + \overline{CS}_j\hat{\delta}_3 + 1*\hat{\delta}_4)$.
[35] We also estimated models separately for elementary and secondary schools (middle/high schools) because there is limited overlap in hiring across schooling levels. Those estimates, which we omit for brevity, show that the borders affect mobility at both schooling levels, but that the effects are larger at the secondary level. In the raw data, leadership hires at the secondary level are more likely to occur across schools and districts, which explains why the border effects are more pronounced there.

sign with the simulation results, and three of the four estimates are statistically significant (the exception being the coefficient in the full-principals model in KC).[36]

Table 7. Pairwise Logit Results. Marginal Effects from Full Models. Dependent Variable: Any Shared Leadership Hire

	Full Model All Leaders	Full Model Full Principals Only
Kansas City (16,836 pairs)		
Pension-Line Crossing	-0.568	-0.148
	(0.280)*	(0.224)
District-Line Crossing	-4.591	-2.473
	(0.980)**	(1.045)*
Free/reduced-lunch similarity	0.008	0.004
	(0.004)*	(0.003)†
Racial similarity (nonwhite)	0.004	-0.000
	(0.005)	(0.003)
Enrollment (log)	1.059	0.163
	(0.144)**	(0.069)*
Distance (miles)	0.063	0.024
	(0.038)†	(0.026)
Distance^2 (miles)	-0.008	-0.003
	(0.003)*	(0.002)
St. Louis (35,245 pairs)		
Pension-Line Crossing	-0.360	-0.190
	(0.108)**	(0.009)*
District-Line Crossing	-2.037	-1.148
	(0.302)**	(0.231)**
Free/reduced-lunch similarity	0.005	0.004
	(0.002)*	(0.002)†
Racial similarity (nonwhite)	0.002	0.003
	(0.002)	(0.002)
Enrollment (log)	0.828	0.357
	(0.077)**	(0.053)**
Distance (miles)	-0.099	-0.090
	(0.024)**	(0.019)**
Distance^2 (miles)	0.004	0.004
	(0.002)*	(0.001)**

** indicates statistical significance at the 1 percent level.
* indicates statistical significance at the 5 percent level.
† indicates statistical significance at the 10 percent level.

[36] The table shows that district borders are important determinants of leadership sharing irrespective of pension borders; the large coefficient on the district-line variable in Table 7 corresponds to the large city-share coefficient in Table 6. One reason that leadership sharing between schools is more common within districts is that leaders are centrally managed and can be reallocated across schools by district administrators. The district-line and city-share coefficients in our models partly reflect this feature of the leadership labor market.

Notes: Robust standard errors are in parentheses. All coefficients and standard errors are multiplied by 100 for ease of presentation. The "principals only" model excludes assistant principals.

We estimate the magnitudes of the border effects implied by the logit models similarly to above – we compute the marginal effect of changing the pension-line variable holding the district-line variable fixed at one and all other variables fixed at their sample averages. Removing the pension border between two schools increases the probability of a shared leadership hire by 85 percent in the KC region, and 96 percent in the STL region. For the full-principals model in STL the probability of a shared hire increases by 81 percent (the corresponding prediction based on the statistically-insignificant estimate in KC is 45 percent).

In the STL region the implied border effects are qualitatively and quantitatively similar using both empirical approaches. In the KC region the simulations imply larger border effects. We briefly investigate the extent to which the discrepancy in the KC region can be explained by the information loss in the pairwise models. Specifically, we re-run the simulations on a modified dataset with the same information loss (the modified dataset allows for, at most, one instance of a shared hire between any two schools). We omit the results for brevity, but the implied border effect in the KC region is 26 percent smaller in the limited-information simulation (moving from 167 to 123 percent). If we place rough confidence intervals around the predictions from each approach using the limited data, we cannot reject that the implied border effects are equal in magnitude.[37]

VII. Policy Implications in Missouri

We now turn to the policy implications for Missouri. From the introductory Figures 1 and 2, we know that the border effects on leadership quality in the urban districts will depend on

[37] We also perform a similar simulation analysis with limited data for STL. The implied border effect declines slightly and remains comparable to the result from the pairwise model (the estimated effect on leadership flows falls from 97 to 88 percent).

there being differences in the applicant pools on different sides of the borders. Are there differences?

We initially investigate this question by examining teachers' average ACT scores.[38] The evidence relating ACT scores to value-added in the classroom is mixed; however, even without direct link to classroom performance, observable differences in teachers' ACT scores across pension systems are likely to imply differences along other dimensions as well.[39] Table 8 shows average ACT scores for incoming teachers between 2005 and 2009 for all PSRS districts, KC, STL, and the geographically close PSRS districts in each region.[40] City teachers clearly have the lowest ACT scores.

Table 8. Average ACT Scores for Incoming Teachers between 2005 and 2009: PSRS, KC, STL and PSRS Neighboring Districts

		Kansas City Region		St. Louis Region	
	PSRS All	KC Schools	Neighbor (PSRS)	STL Schools	Neighbor (PSRS)
Average ACT Score	22.83	21.26	23.00	21.96	23.40
	(3.87)[a,b]	(4.69)	(3.74)[a]	(4.72)	(4.11)[b]
N (teachers)	11,501	177	1190	336	1304
Avg. % Free/Reduced Lunch	31.7	59.9	23.9	70.8	29.6
Avg. % Disadv. Minority	7.1	70.9	27.1	83.3	45.0
Districts	539	1	11	1	21

Notes: The neighboring districts are within a commutable distance to the city in each region. ACT scores are reported for all teachers for whom they are observed, which is roughly half of all of the new teachers who are identified in the statewide data panel. Some institutions that train teachers in Missouri do not require the ACT. Standard deviations are in parentheses. Superscripts indicate statistically significant differences ($p < 0.01$) from Kansas City (a) and St. Louis (b), respectively (test results are reported within region and for each city district relative to PSRS).

[38] The ACT data are available in a separate higher-education database in Missouri. ACT scores are reported for all individuals for whom they are available. Missing scores are fairly common. One reason is that many teacher preparation programs in Missouri do not require the ACT.

[39] Harris and Sass (forthcoming) find that teachers' college entrance exams scores are not related to productivity in the classroom as measured by value-added, which suggests a limitation of using ACT scores to compare teachers. However, Ferguson and Ladd (1996) find a positive relationship using school-aggregated data.

[40] The neighboring PSRS districts in Tables 8 and 9 are the same districts as in Table 3, broken out by region.

In Table 9 we turn to credentials for observed school leaders. First, we compare leaders using their scores on the School Leaders Licensure Assessment (SLLA), a standardized examination administered by the Educational Testing Service that is required of all applicants for school administration licenses in Missouri. The SLLA is designed to measure applicants' leadership capacities as defined by a set of national standards for leadership practice (Reese and Tannenbaum, 1999). The average score for school leaders in Missouri is roughly 180, and the standard deviation is 7.5.

Table 9. Licensure Exam Scores and College Quality for School Leaders: PSRS, KC, STL and PSRS Neighboring Districts

		Kansas City Region		St. Louis Region	
	PSRS All	KC Schools	Neighbor (PSRS)	STL Schools	Neighbor (PSRS)
Licensure Exam Scores					
Average Score	178.2	172.8	179.8	175.1	178.8
	(7.4)[a,b]	(7.2)	(7.0)[a]	(7.1)	(7.3)[b]
N (leaders)	4,099	163	261	222	322
College Quality: All					
High Quality	0.174	0.112	0.184	0.077	0.233
	(0.379)[a,b]	(0.316)	(0.388)[a]	(0.267)	(0.423)[b]
N (leaders)	8,873	339	803	377	1101
College Quality: MO Specific					
High Quality	0.187	0.083	0.177	0.078	0.262
	(0.390)[a,b]	(0.276)	(0.382)[a]	(0.269)	(0.440)[b]
Low Quality (Public)	0.064	0.223	0.056	0.566	0.162
	(0.245)[a,b]	(0.417)	(0.229)[a]	(0.496)	(0.369)[b]
N (leaders)	7,089	193	575	295	809
Avg. % Free/Reduced Lunch	31.7	59.9	23.9	70.8	29.6
Avg. % Disadv. Minority	7.1	70.9	27.1	83.3	45.0
Districts	539	1	11	1	21

Notes: The neighboring districts are within a commutable distance to the city in each region. Licensure exam scores are available for school leaders from 2000-2009. College quality is available throughout the data panel and is coded based on the institutions where leaders obtained their initial bachelor's degrees. Standard deviations are in parentheses. Superscripts indicate statistically significant differences ($p < 0.01$) from Kansas City (a) and St. Louis (b), respectively (test results are reported within region and for each city district relative to PSRS).

We also use a supplementary data file from the Department of Elementary and Secondary Education to measure undergraduate-institution quality for school leaders. We consider Missouri graduates exclusively, who make up roughly 75 percent of the leaders in the data, and all graduates. We identify four "high-quality" universities in Missouri: Truman State University in Kirksville, the University of Missouri in Columbia, the University of Missouri in Rolla, and Washington University in St. Louis.[41] We also identify the three public institutions in Missouri with the lowest entrance exam scores: Harris-Stowe State University in St. Louis, Lincoln University in Jefferson City, and Missouri Western State University in St. Joseph.[42] For the larger, nationwide sample of graduates we use the following (admittedly rough) guideline to categorize "high-quality" universities: if a university has an average entrance-exam score that is at least as high as at the University of Missouri Columbia, or if it is the flagship university in its state, we code it as a "high quality" university. For the national sample we do not code low-quality universities.[43]

Table 9 shows that city leaders score lower on the SLLA exam and come from lower-quality universities relative to leaders on the outside. The differences across the pension systems are large along all dimensions, with a particularly notable difference being that city leaders are

[41] We hesitate to use the word "selective" because in the broader sense these universities may not be considered to be selective. However, within Missouri, they are clearly the most prestigious universities for undergraduates. For example, Truman State University and the University of Missouri at Rolla have higher average entrance exam scores than the University of Missouri, Columbia, which is the state's flagship research university.

[42] We do not categorize the bottom tail of the distribution of private universities because there are many small universities and we do not have the data to distinguish them accurately. For the public schools we selected these three universities because there is a large difference in average ACT scores between them and the other public universities. For example, Missouri Western has the highest average ACT score of the three, but it is still a half of a standard deviation below the average ACT score for the next-lowest-ranked public university. The average ACT scores at Lincoln University and Harris-Stowe State are another half-standard deviation below Missouri Western.

[43] Like with teachers' ACT scores, our measures of leadership quality are imperfect. In fact, Clark et al. (2009) find that the selectivity of a principal's undergraduate institution, as measured by median SAT scores, is not related to principal quality in New York City. However, while Clark et al. (2009) do not find a general relationship, their analysis does not rule out the possibility that graduates from particularly unselective colleges perform worse. A large fraction of leaders in the border regions in Missouri, and in the city districts in particular, are graduates from such colleges (see Table 9).

much more likely to have attended one of the least selective public universities in Missouri. The evidence in Table 9 is descriptive, and therefore merely suggestive, but it is consistent with the pension-induced rigidities in the labor market adversely affecting leadership quality in the city districts. While the pension borders may not be the only reason for differences in leadership quality on different sides of the borders, they are likely to be a contributing factor. Moreover, it is important to recognize that even if the pension borders play little or no role in creating the initial quality differences, by imposing a large tariff on educators who cross boundaries, they represent a major impediment to eliminating the gaps.

VIII. Discussion and Conclusion

Leadership quality is increasingly recognized as a key contributor to student success in schools (Brewer, 1993; Cannon et al., 2010; Clark et al., 2009; Coelli and Green, 2010; Dhuey and Smith, 2010; Grissom, 2011; Grissom and Loeb, 2011). We examine an important and previously unaddressed feature of the labor market for school leaders – namely, school leaders, and prospective school leaders, are heavily invested in defined-benefit pension plans that restrict mobility. We show that pension borders introduce substantial rigidities into the leadership labor market, resulting in the misallocation of labor to school-leadership positions. When a pension border divides the labor market for a small school district, or when there is an underlying quality difference in the applicant pools on different sides of a pension border, leadership quality will be reduced. Furthermore, the quality of leader/school matches will be lower near pension borders. Applying recent evidence from Jackson (2010) to the leadership context suggests that the lower match quality will adversely impact students in K-12 schools.[44]

[44] And, of course, because the pension borders will also restrict teacher mobility, teacher/school match quality will be lower in border regions.

Missouri offers a unique opportunity to examine pension-border effects because educators belong to three different pension systems and there is not reciprocity between them. That is, the three systems within Missouri are separated in the same way that state-level pension systems across the country are separated. Our findings are of interest nationally given that 22,000 miles of state lines across the country serve as pension boundaries.

We conclude by noting that the issues raised by our analysis are, in principle, straightforward to resolve. One solution would be for state governments to compensate school leaders for the costs of crossing pension system lines to reduce inefficiencies resulting from the labor misallocation. Again noting the recurring finding that high-quality school leaders contribute greatly to student success, the benefits of such a policy may outweigh the costs in some instances. Furthermore, in Missouri, there would be an equity gain in addition to the efficiency gain because the urban districts educate disproportionately poor and minority students. The other option, which is conceptually straightforward but would be politically challenging to implement, is to remove the pension borders. This can be achieved by integrating pension systems nationwide, or by moving educators into retirement plans that do not discourage mobility (e.g., defined-contribution or cash balance plans). It is interesting to contrast the absence of interstate reciprocity and collaboration in teacher pension plans with the growing reciprocity and efficiency that has developed in teacher licensing, largely as a result of collaboration by certification staff in the various states.[45] Ironically, while mobility restrictions due to educator licensing barriers have declined, the pension-wealth penalties associated with

[45] Specifically, the National Association of State Directors of Teacher Education and Certification (NASDTEC) Interstate Agreement is designed to reduce or eliminate barriers to movement of certified teachers and administrators between states and jurisdictions. The teacher agreement has been signed by 48 states and the administrator agreement has been signed by 32 states (http://www.nasdtec.org/agreement.php). The National Board for Professional Teaching Standards (NBPTS) also seeks to have its advanced certificates for teachers and principals recognized as a valid license for all fifty states. In contrast, there is little or no effort to create similar reciprocity for teacher pension plans. See Ruppert (2001) and Gates (1996).

interstate educator mobility have grown over recent decades due to benefit enhancements in the state plans.

References

Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak and Christopher R. Walters. 2010. Who Benefits from KIPP? NBER Working Paper No. 15740.

Bartelsman, Eric J. and Mark Doms. 2000. Understanding Productivity: Lessons from Longitudinal Microdata. *Journal of Economic Literature*, 38(3), pp. 569-594.

Brewer, Dominic J. 1993. Principals and Student Outcomes: Evidence from U.S. High Schools. *Economics of Education Review* 12(4), 281-292.

Bureau of Labor Statistics. 2011. Labor Market Areas, 2011. Report available at: http://www.bls.gov/lau/lmadir.pdf

Cameron, A. C., J. B. Gelbach, and D. L. Miller. (2011). Robust inference with multi-way clustering. *Journal of Business and Economics Statistics* 29(2), pp. 238-249.

Cannon, Sarah, David Figlio and Tim Sass. 2010. The Policy Choices of Effective Principals. Unpublished manuscript. Northwestern University.

Clark, Damon, Paco Martorell and Jonah Rockoff. 2009. School Principals and School Performance. CALDER Working Paper No. 38.

Clark, Robert and Craig, Lee A. (2011). "Determinants of the Generosity of Pension Plans for Public School Teachers, 1982-2006." *Journal of Pension Economics and Finance*,10(1), 99-118.

Coelli, Michael and David Green. 2010. Leadership Effects: School Principals and Student Outcomes. Unpublished manuscript. University of Melbourne.

Costrell, Robert M. and Michael Podgursky. 2009. Peaks, Cliffs, and Valleys: The Peculiar Incentives in Teacher Retirement Systems and Their Consequences for School Staffing. *Education Finance and Policy* 4(2), 175-211.

Coile, Courtney and Jonathan Gruber. 2007. Future Social Security Entitlements and the Retirement Decision. *The Review of Economics and Statistics* 89(2), pp 234-246.

Cullen, Julie Berry and Michael J. Mazzeo. 2008. Implicit performance awards: an empirical analysis of the labor market for public school principals. Unpublished manuscript. University of California, San Diego.

Dhuey, Elizabeth and Justin Smith. 2010. How Important Are School Principals in the Production of Student Achievement? *Unpublished manuscript*.

Ferguson, Ronald F. and Ladd, Helen F. 1996. "How and Why Money Matters: An Analysis of Alabama Schools," in *Holding Schools Accountable: Performance Based Reform in Education*, ed. Helen F. Ladd. Washington: Brookings Institution.

Gates, Gary I. 1996. *Pension Portability for State and Local Government.* New York: Millbank Memorial Fund.

Grissom, Jason A. (forthcoming). Can Good Principals Keep Teachers in Disadvantaged Schools? Linking Principal Effectiveness to Teacher Satisfaction and Turnover in Hard-to-Staff Environments. *Teachers College Record*.

Grissom, Jason A., and Susanna Loeb. 2011. Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills. *American Educational Research Journal*, in press.

Gustman, Alan L. and Thomas L. Steinmeier. 1989. An Analysis of Pension Benefit Formulas, Pension Wealth and Incentives from Pensions. *Research in Labor Economics* 10, pp 53-106.

-- 2001. What People Don't Know About Their Pensions and Social Security: An Analysis Using Linked Data From the Health and Retirement Study. In William G. Gale, John B. Shoven and Mark J. Warshawsky, editors, *Public Policies and Private Pensions*. Washington, D.C.: Brookings Institution.

Harris, Douglas N. and Tim R. Sass (forthcoming). Teacher Training, Teacher Quality and Student Achievement. *Journal of Public Economics*.

Jackson, Kirabo. 2010. Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. NBER Working Paper No. 15990. Cambridge, MA.

Jacob Brian. 2007. The Challenges of Staffing Urban Schools with Effective Teachers. *The Future of Children* 17(1), pp. 129-153.

Loeb, Susanna, Demetra Kalogrides, and Eileen Horng. 2010. Principal Preferences and the Uneven Distribution of Principals across Schools. *Education Evaluation and Policy Analysis*, 32(2), pp. 205-229.

Loeb, Susanna and Luke Miller. 2006. A Review of State Teacher Policies: What are they, what are their effects, and what are their implications for school finance? Policy Analysis for California Education. University of California – Berkeley.

Mitchell, Olivia, David McCarthy, Stanley Wisniewski, Paul Zorn. 2001. "Developments in State and Local Pension Plans." In O. Mitchell and E. Hustead. (eds) *Pensions and the Public Sector*. Philadelphia: University of Pennsylvania Press, 11-40.

Mortensen, Dale T. and Christopher A. Pissarides. 1999. New developments in models of search in the labor market. *Handbook of Labor Economics*, vol. 3B, pp. 2567-2627.

Munnell, Alicia. 2000. The Impact of Mandatory Social Security Coverage of State and Local Workers: A Multi-State Review. AARP Public Policy Institute.

Myung, Jeannie, Susanna Loeb and Eileen Lai Horng (forthcoming). Tapping the Principal Pipeline: Identifying Talent for Future School Leadership in the Absence of Formal Succession Management Programs. *Education Administration Quarterly*.

Podgursky, Michael and Mark Ehlert. 2007. Teacher Pensions and Retirement Behavior. CALDER Working Paper No. 5.

Reese, Clyde M. and Richard J. Tannenbaum. 1999. Gathering Content-Related Validity Evidence for the School Leaders Licensure Assessment. *Journal of Personnel Evaluation in Education* 13(3), pp. 263-282.

Ruppert, Sandra S. 2001. *Improving Pension Portability for K-12 Teachers.* Washington DC: State Higher Education Executive Officers (SHEEO) (February).

Stock, James H. and David A. Wise. 1990a. Pensions, the Option Value of Work, and Retirement. *Econometrica* 58(5), pp 1151-1180.

-- 1990b. The Pension Inducement to Retire: An Option Value Analysis. In David A. Wise, (Ed.) *Issues in the Economics of Aging* (Chicago: University of Chicago Press).

Thompson, Samuel B. (2010). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 19(1), pp. 1-10.

Appendix A
Details for the Pension-Wealth Calculations

The pension-wealth calculations are based on the rules for each system in the year-2000 (see Appendix Table C.1 – the mobility penalties are large regardless of which set of rules we use). We use a real interest rate of four percent.[46] The choice of the discount rate affects our estimates of the absolute mobility costs; therefore, we also measure mobility costs using the percentage change in pension wealth relative to the baseline case where no across-system move occurs. While our estimates of the absolute mobility costs are sensitive to the discount rate, the percent-change estimates are not (for reasonable adjustments). The mobility-cost calculations are reported in Appendix Table C.2 for moves in all possible directions in Missouri and at three different ages for the representative teacher (age-38, age-45, and age-50).

For our calculations we project the representative teacher's survival probabilities and wages, and some parameters of the Social Security system. The survival probabilities come from the Cohort Life Tables provided by the Social Security Administration. We use survival probabilities for women, and note that expected mobility costs (in absolute dollars) are slightly lower for men because they have shorter expected life spans. We project wages under the different career scenarios based on our analysis of the wage-growth profiles for all teachers and leaders in Missouri over the course of our data panel. The remainder of this appendix describes the wage projections.

To project wages we begin with the sample-average wage among teachers in Missouri who are the same age as the mover in the year prior to the potential leadership move. So, for example, for the age-38 mover we begin with a teacher earning the average wage in the state for age-37 teachers. This wage is projected forward and backward to create a full wage profile over the career cycle under the different career scenarios. Projecting wages forward would suffice for calculating system pension wealth because *FAS* depends only on the final few years of earnings; however, Social Security wealth depends on a long history of earnings (up to 35 years). Therefore, we also project wages backward for the calculations involving the city districts.

For teaching years we project wages forward and backward using the parameters from an OLS regression where we regress earnings on a cubic function of experience for all teachers in the data panel (1994-1995 and later – in total, we use data from over 130,000 teachers to estimate this regression). For moves into leadership we make two projections: we project the immediate wage increase corresponding to the move, and subsequent earnings growth. On average in the data, moves to leadership correspond to a 30-percent wage increase. The prediction function for the immediate wage increase is again a cubic of teaching experience. Because we combine assistant principals and principals as "school leaders", the wage-increase function averages all moves from teaching to either position. However, the results are qualitatively similar if we focus only on full principals. This is because the wage increase corresponding to a move from teaching to an assistant principal position is much larger than the

[46] This discount rate falls somewhere in the middle of what others in the literature have used. For example, Coile and Gruber (2004) use a real discount rate of 6 percent, and Costrell and Podgursky (2009) use a real discount rate of 2.5 percent.

wage increase corresponding to a move from an assistant- to full-principal position. A large share of the salary increase is associated with moving from a 9-month to 11-month contract.

Once a switch occurs, the wage-growth function for school leaders differs from that of teachers. For one, it depends on leadership experience, not teaching experience.[47] Also, wages for school leaders do not grow smoothly like teacher wages. Therefore, instead of projecting out leadership wage growth using a cubic experience function, we use the data to estimate a step function that measures the average gain in salary going from one to two years of leadership experience, two to three years, three to four years, etc. The step function is estimated using the observed wage profiles of school leaders who entered leadership during the 1994-1995 school year or later (so that we could calculate leadership experience). The real return to experience is highest early in leaders' careers, and flattens out at around two percent per year by year nine.[48] Also, because school leaders who remain in leadership longer are likely to be non-randomly selected (much more so than teachers), we include individual fixed effects in the leadership wage-growth function.[49] Finally, for both teachers and principals we allow wages to grow with expected inflation.[50]

[47] That is, empirically, teaching experience does not predict principal wage growth conditional on principal experience. This is consistent with evidence from Clark et al. (2009), who show that principal effectiveness is related to prior leadership experience, but not to prior teaching experience.
[48] We allow the final step to repeat indefinitely for principal spells lasting more than 9 years.
[49] For consistency we also estimated teacher-wage-growth models with individual fixed effects, but as expected, the individual fixed effects in the teacher projection function are less important. We have considered the sensitivity of our findings to using a model without fixed effects to determine the projection parameters for teacher wages, and this does not qualitatively affect our results.
[50] The wage-growth projection functions predict real wage growth.

Appendix B
Procedure for Determining the Effective Number of Simulations

We resample from the same underlying pool of schools to construct the artificial districts in each region. Therefore, in addition to the data being potentially correlated across artificial districts within each simulation, the data are also correlated across simulations. Two-way clustering approaches (Thompson, 2010; Cameron et al., 2011) cannot be used because the artificial districts within each simulation are unordered, and there are not cluster labels. This makes it inadvisable, and undesirable, to impose assumptions on the variance-covariance structure across simulations (although, as noted in the text, we do allow for within-simulation error covariance). To deal with the across-simulation correlation in the data, we develop a procedure to estimate the effective simulation counts given the actual number of simulations that we run in each model (1,000). The effective simulation calculations are based on an *ex post* analysis of the degree of similarity of error terms from the regression in (3) across simulations.

We compare the 1,000 batches of residuals from the regression (normalized to unit variance), where each batch contains the 10 residuals corresponding to the 10 artificial districts, to analogous batches of residuals from a Monte Carlo. First, we sort the 10 regression residuals in each batch in ascending order and compute the variance of the *k*th largest term, s_k^2 (k=1,..,10), across the 1,000 batches. Then we conduct a Monte Carlo where we draw 1,000 batches of 10 errors from q independent batches, where the Monte Carlo batches come from a comparable distribution to the distribution of errors in the simulated data.[51] We control q and draw 1,000 batches at each level of q from 1 to 1,000. As with the residuals from regression (3), we sort the 10 errors from each Monte Carlo batch in ascending order and compute the variance of the *k*th (k=1,..,10) largest term across the 1,000 batches. We repeat the procedure 10,000 times and compute the average variance σ_k^2 (k=1,..,10). σ_k^2 is increasing in the number of independent batches, q. In the extreme case with q =1, all 1,000 batches are identical and $\sigma_k^2 = 0$ (for k=1,..,10). If the errors in each batch were iid, then as q increases σ_k^2 approaches the variance of the k-order statistics.[52]

By controlling q, it is straightforward to track the amount of underlying independent data in the Monte Carlo. We estimate the effective number of simulations in each model, q, by equating the cross-simulation variances of the ordered residuals, s_k^2, and the Monte Carlo variances, σ_k^2, for (k=1,..,10). We take the implied q averaged over *k*. The resultant *q's* for the sample range from 350 to 450. They are consistently larger in the STL region; and within region, in the full-principal models.[53]

---

[51] The 10 errors in each batch in the Monte Carlo are drawn from a joint normal distribution with unit variance and a given within-batch correlation, which reflects the estimated within-batch correlation in the errors in the simulated data (e.g. -0.03 for schools bordering Kansas City).
[52] The errors within each batch are not iid because of the within-batch correlation. We also conduct simulations based on draws that are iid within batches. The estimated effective simulation counts do not change substantially.
[53] Sample programs for the Monte Carlo are available from the authors upon request.

Appendix C
Supplementary Tables

Appendix Table C.1. Changes to Key Parameters of Pension Systems, 1995 – 2009 (there were no changes after 2002). Initial Parameters as of 1995 are Reported in Row 1.

	PSRS	KC	STL
1995	Formula factor 0.023, early retirement by 55-25 rule, COLA cap 65 percent	Formula factor 0.0175, Rule of 75, no COLA	Formula factor 0.0125, Rule of 85, no COLA
1996	Implement unrestricted "25 and out"		
1997	COLA cap increased from 65 to 75 percent		
1998			
1999	Formula factor raised to 0.025 for full retirement (with corresponding upward adjustments for early retirement)		
2000	Implement Rule of 80 FAS changed to highest three years of salary	Formula factor increased to 0.020	Formula factor increased to 0.020, *F*YOS* capped at 0.60 (previously uncapped)
2001	COLA cap increased to 80 percent		Implement DROP provision
2002	Formula factor increased to 0.0255 if YOS $\geq$ 31 (new factor applies to *all service years* for eligible individuals)		

* The DROP provision in St. Louis allows teachers and principals who qualify for full retirement to delay retirement for up to four years without losing their pension payments for that time. Pension wealth is frozen, and those payments are "dropped" into an account that is paid out at the time of work stoppage.

Appendix Table C.2. Pension-Wealth Costs of Across System Moves into Leadership Positions Based on Peak Value. Year-2000 Pension Rules. Year-2009 Dollars. Women.

	Constrained Retirement Date in Second System		Work Until Age 65 in Second System	
	Comparison Between Across-System and Within-System Moves to Leadership	Comparison of Within-System Career-Teacher Scenario and Across-System Move to Leadership	Comparison Between Across-System and Within-System Moves to Leadership	Retirement Age Without/With Move (Determined by Peak value)
PSRS to KC				
Move at: 38	-283,638 (-63)	-156,093 (-49)	-171,978 (-38)	53 / 65
45	-339,682 (-65)	-204,658 (-52)	-216,116 (-41)	53 / 65
50	-85,868 (-14)	-52,486 (-11)*	3,987 (+1)	56 / 65
PSRS to STL				
Move at: 38	-312,871 (-70)	-185,326 (-58)	-168,081 (-38)	53 / 65
45	-358,815 (-68)	-223,791 (-57)	-212,680 (-40)	53 / 65
50	-132,436 (-22)	-52,486 (-11)*	6,738 (+1)	56 / 65
KC to PSRS				
Move at: 38	-151,768 (-44)	-63,522 (-25)	-39,766 (-12)	54 / 65
45	-164,237 (-40)	-73,922 (-23)	-37,570 (-9)	54 / 65
50	-53,453 (-12)	-30,033 (-8)*	89,136 (+19)	54 / 65
KC to STL				
Move at: 38	-181,190 (-53)	-92,944 (-36)	-72,766 (-21)	54 / 65
45	-179,182 (-44)	-88,867 (-28)	-68,847 (-17)	54 / 65
50	-66,138 (-14)	-30,033 (-8)*	45,598 (+10)	54 / 65
STL to PSRS				
Move at: 38	-141,445 (-41)	-51,310 (-20)	-46,003 (-13)	55 / 65
45	-198,204 (-48)	-103,736 (-33)	-87,382 (-21)	55 / 65
50	-183,799 (-39)	-96,915 (-25)	-56,970 (-12)	55 / 65
STL to KC				
Move at: 38	-142,957 (-42)	-52,822 (-21)	-82,901 (-24)	55 / 65
45	-193,231 (-47)	-98,763 (-32)	-122,095 (-30)	55 / 65
50	-187,890 (-40)	-101,006 (-27)	-105,065 (-22)	55 / 65

Notes: Percentages of baseline pension wealth under the relevant no-move scenario are reported in parenthesis. An '*' indicates that under the constrained retirement scenario the individual will not become vested in the second pension system, which occurs in several cases when we use the career-teacher scenario as the baseline.

Appendix Table C.3. Within- and Out-of-District Leadership Hires for PSRS, KC, STL, and Subsamples of PSRS Districts. Various Subsamples of School Leaders.

	PSRS All	KC	STL	PSRS Neighbors	PSRS Disadvantaged	PSRS Large Districts
First-Time Leadership Hires						
Within-District Hires	60.3	89.9	91.4	67.7	60.1	76.9
Out-of-District Hires	36.2	6.3	6.8	29.7	37.4	20.7
Unknown (in-state)	3.5	3.9	1.8	2.6	2.5	2.4
N	5023	207	278	1028	203	801
First-Time Leadership Hires Direct from Teaching						
Within-District Hires	60.6	92.8	91.2	67.8	54.2	78.7
Out-of-District Hires	39.4	7.2	8.8	32.2	45.8	21.3
Unknown (in-state)	-	-	-	-	-	-
N	3421	138	147	627	118	507
Full Principals Only, All Hires						
Within-District Hires	53.4	86.5	87.2	64.4	64.6	77.5
Out-of-District Hires	44.5	12.8	12.6	34.4	34.1	21.8
Unknown (in-state)	2.0	0.7	0.3	1.2	1.2	0.7
N	6342	282	366	1086	246	843
Age 45 or Younger (All Hires)						
Within-District Hires	54.4	88.1	86.7	59.8	56.7	70.5
Out-of-District Hires	44.6	10.1	12.2	39.7	43.3	29.0
Unknown (in-state)	1.0	1.8	1.1	0.6	0.0	0.5
N	6327	218	180	1382	217	1205
Avg. % Free/Reduced Lunch	31.7	59.9	70.8	27.6	60.3	13.5
Avg. % Disadv. Minority	7.1	70.9	83.3	38.9	79.3	18.2
Average Enrollment	1495	32839	39948	6708	3121	17948
Number of Districts	539	1	1	32	9	10

Notes: PSRS Neighbors include districts that are geographically adjacent to either KC or STL, or are adjacent to the adjacent districts. The neighboring districts are within a commutable distance to the city in each region.

Appendix Table C.1

Appendix Table C.1 shows changes to the pension-system rules over the course of the data panel. Benefits have become more generous over time in each system. The calculations corresponding to Figure 3 in the text are based on the system rules as of the year-2000. Mobility costs have risen modestly as the benefits within each system have improved.

Appendix Table C.2

Appendix Table C.2 shows the costs of pension-border crossings in Missouri for the representative teacher moving into leadership at three ages: 38, 45 and 50. These ages correspond to the 50th, 75th and 90th percentiles of the unconditional distribution of age-at-entry for new school leaders in Missouri who come directly from teaching (for whom the simulations are relevant based on the wage projections). The pension-accrual profiles in Figure 3 correspond to the age-38 move scenarios for the PSRS teacher.

The table compares pension-wealth accrual under the scenario where an across-system move into leadership occurs to scenarios where (1) the move into leadership occurs without crossing a pension border, and (2) no leadership promotion occurs and the individual remains in teaching within the same pension system. In the first two columns of the table we constrain retirement in the second system to occur at the same age as in the first system based on our peak value calculations. In column (3) we allow retirement in the second system occur at the age of 65, which is late for educators (Costrell and Podgursky, 2009).

There is some variability in the cost of the move depending on its direction, although the costs in most cases are high. One source or variability in the mobility costs is that it is more costly to move from the more-generous PSRS system, which further disadvantages the city districts. Also, differences in the early-retirement provisions across systems are important – it is more costly to leave the plans that have more-generous early retirement provisions for younger workers, and vice-versa for older workers.[54] On a related note, late-career leaders in PSRS and KC stand to lose less if they move, and in some cases can marginally gain pension wealth because they can double-dip. That is, they can begin collecting their first pension while working in the new system. However, Appendix Table C.3 shows that double dipping cannot explain what little cross-border mobility we see in the data.

Appendix Table C.3

Appendix Table C.3 shows within- and out-of-district hire rates for subsamples of school leaders in Missouri. The mobility rates within each subsample are generally similar to what is reported for all leaders in Table 3 in the text.

On the issue of double-dipping, note that the out-of-district hire rate is not lower in the city districts for movers who are aged 45 or younger. Based on our calculations in Appendix

[54] For example, consider the age-50 mover, for whom the constrained-retirement costs of exiting KC and PSRS are much smaller than the costs of exiting STL. This is because in both PSRS and KC the mover can take advantage of an early-retirement provision and begin collecting benefits immediately (25-and-out and the rule-of-75, respectively). In STL, the move is more costly because the age-50 mover is still 10 years away from the rule-of-85 at the time of the move. Many studies in the general retirement literature focus on replacement rates as a measure of pension generosity (e.g., Mitchell, et.al , 2011; Loeb and Miller, 2006; Clark and Lee, 2011). Appendix Table C.2 shows that the age when individuals can begin collecting is also important.

Table C.2, one might expect lower mobility rates for this group if across-system mobility is primarily driven by older movers, who can draw on their first pension while working in the new system. But the data are not consistent with double-dipping playing an important role in the labor market. We offer two explanations for this result. First, it could be that late-career educators – who would be eligible to double-dip – are more likely to turnover after shorter leadership spells, which would discourage demand. Second, on the supply side, despite the potential for pecuniary gains associated with late-career switches across borders, individuals' marginal returns to labor and leisure during the late stage in their careers may be such that the move is undesirable. This explanation seems more plausible given that even without double dipping, educators who are eligible for full retirement in Missouri already collect generous pensions.